POLICING VIDEO CONTENT ON YOUTUBE, FACEBOOK AND TWITTER

Platforms’ New Efforts to Block Offensive Clips Explained

JUNE 2018
Paul Verna
Contributors: Annalise Clayton, Sean Creamer
YouTube, Facebook and Twitter have faced a deluge of marketer objections over offensive videos and other problematic content on their platforms. In the past year alone, these services have found themselves at the center of controversies over everything from live streamed suicides to videos that promote terrorism to racist tweets from celebrities and others.

These incidents—combined with a backlash over a range of related issues including promoting fake news, providing a forum for abusive trolls, influencing the 2016 US presidential election and misusing customer data—have put the platforms on the defensive.

- Over the past few months, each platform has taken steps to beef up security and promote brand safety, particularly around video. These efforts have involved a combination of human monitors and machine learning analytics—with an emphasis on the latter, given the scope of the issue.

- The companies have also taken steps to provide more transparency to advertisers, regulators, users and anyone else wanting to know how the platforms deal with troublesome content.

- There are indications that these companies are making progress in beating back the tide of offensive content through improved algorithms and a sharp focus on the issue, but most experts say no platform will ever be 100% safe as long as it traffics in user-generated media.

- Despite ad boycotts, public outcries, regulatory scrutiny and punishing media coverage, the brand safety issues that have plagued these platforms have not had an apparent effect on their bottom lines or share prices so far. Marketers continue to see value in advertising on these platforms.

**WHAT'S IN THIS REPORT?** This report explores how YouTube, Facebook and Twitter are policing user-uploaded video to block content that violates their guidelines. These platforms have recently stepped up their efforts to promote brand safety in the wake of scandals relating to offensive content and data privacy, which in turn led to ad pullbacks.

**CONTENTS**

2 Policing Video Content on YouTube, Facebook and Twitter: Platforms’ New Efforts to Block Offensive Clips Explained

4 Brand Safety Under Threat

4 YouTube

9 Facebook

12 Twitter

14 Brand Safety: A Shared Responsibility

14 Key Takeaways for Marketers

15 eMarketer Interviews

16 Read Next

16 Sources

16 Editorial and Production Contributors
2,400 MINUTES

Amount of video uploaded every minute of “real” time*

Combined running time of videos watched every day**

144,155 YEARS

70 million HOURS

Total watch time of videos created in YouTube Spaces**

Views of “Despacito” on YouTube**

5.24 BILLION

KEY STAT

The amount of video that passes through YouTube on any given day is staggering. If a day’s clips were watched back to back, it would take 150 millennia (not millennials!) to go through them.

Sources: * The Wall Street Journal, “YouTube Tops 1 Billion Hours of Video a Day, on Pace to Eclipse TV,” Feb 27, 2017; ** YouTube; “Despacito” view count as of June 19, 2018; eMarketer calculations
In early 2017, YouTube found itself grappling with a flurry of issues that threatened its image as a brand-safe haven for video and display advertising. In February, PewDiePie—at the time the YouTuber with the most subscribers—posted anti-Semitic videos on his channel. A month later, YouTube inadvertently ran ads from major companies next to videos supporting terrorism.

YouTube responded quickly to these crises by severing ties with PewDiePie, removing the offensive videos and apologizing to marketers. However, these measures did not stop major brands, including GlaxoSmithKline, Johnson & Johnson, Enterprise, AT&T and Verizon, from pulling their ads from the platform, at least temporarily. The backlash gathered steam throughout the first quarter of last year, and soon hundreds of companies were pulling back from YouTube. In late March 2017, an analyst at Nomura Instinet estimated that the boycott would cost Google up to $750 million.

Over the course of the year, YouTube continued to take steps to prevent brand safety lapses, and the issue died down. By October, a Credit Suisse report noted that most advertisers had returned to YouTube, and that the issues had not caused Google to lose any business.

Other brand safety issues followed, however, and again they were met by boycotts or threats from advertisers. At its May 2018 Digital Content NewFronts presentation to advertisers, YouTube again found itself on the defensive, as it had a year earlier. In Q2 2018, the company released its first-ever Community Guidelines Enforcement report, designed to provide transparency into its video policing efforts and reassure advertisers that it is working hard to avert further failures.

During the same time span, Facebook and Twitter also grappled with their share of issues around data security and brand safety. Starting in late 2016, both companies were heavily scrutinized for their purported roles in spreading fake news and potentially influencing the US presidential election. Then, in early 2018, Facebook was rocked by the Cambridge Analytica scandal, in which the UK-based data firm with ties to the Trump campaign was found to have misused data from as many as 87 million Facebook users in the context of the 2016 election.

The core issues that these platforms have faced are different, but they share something in common: They threaten the image the companies have tried to cultivate as spaces in which average citizens can speak their minds, and marketers can reach vast numbers of people.

There’s another common thread throughout these brand safety issues. Despite actual and threatened boycotts, public outcries, testimony before the US Congress and UK Parliament, and a general feeling of unease with how these platforms have managed—or even accepted—their roles in the digital media ecosystem, their businesses have flourished while they faced serious existential threats. There’s no guarantee that business success will continue, but if the past 15 months are any indication, YouTube, Facebook and Twitter have proven, if nothing else, that they are resilient platforms that deliver value to users and advertisers alike.

Addressing a capacity crowd at YouTube’s annual Brandcast event on May 3, 2018, CEO Susan Wojcicki acknowledged the company’s struggles with brand safety over the past year.

“With openness also come challenges, as some have tried to take advantage of our services,” she said. “There isn’t a playbook for how open platforms operate at our scale. But the way I think about it is, it’s critical that we are on the right side of history.”

It was the second year in a row Wojcicki was compelled to bring up a topic she’d likely rather avoid. In May 2017, after some brands had started boycotting YouTube over the inadvertent placement of their ads next to offensive videos, Wojcicki apologized to Brandcast attendees.

“The last couple weeks have been challenging for some of you,” she said at the event, which is designed to showcase content for advertisers. “I want you to know that we have taken your feedback to heart. We work hard every day to earn our advertisers’ and agencies’ trust. We apologize for letting some of you down. We can, and we will, do better.”

The resurgence of the brand safety issue belies the steps YouTube has taken to shield its platform against content that violates its terms of service. Over the past year, the company has hired as many as 10,000 staffers to monitor content, and invested in artificial intelligence (AI)
technology to identify and block potentially problematic videos. It has also lived up to its pledge to be more transparent, launching the first of a series of reports in April 2018 that shed light on its policing efforts.

YouTube removed 8.3 million videos in Q4 2017, according to the YouTube Community Guidelines Enforcement report, which the company said it would issue quarterly starting in April. More than 80% of those videos were flagged by the platform’s automated systems, while 13.7% were flagged by an “individual trusted flagger.” Average users accounted for just under 5% of flags, and NGOs and government agencies made up less than 1% each.

<table>
<thead>
<tr>
<th>Number of YouTube Videos Removed Worldwide, by Flagger Type, Q4 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automated flagging</td>
</tr>
<tr>
<td>Individual trusted flagger</td>
</tr>
<tr>
<td>User</td>
</tr>
<tr>
<td>NGO</td>
</tr>
<tr>
<td>Government agency</td>
</tr>
</tbody>
</table>

Note: total videos removed=8.3 million
Source: YouTube, “YouTube Community Guidelines Enforcement,” April 23, 2018

Of the videos blocked by humans, the largest percentage were flagged for sexual content, followed by spam or misleading content. Other videos were flagged because they were considered hateful or violent, or because they depicted harmful or dangerous acts, or child abuse. Videos that promote terrorism made up only 1.6% of flags, despite the disproportionate attention this category received in some of YouTube’s most publicized brand safety issues.

<table>
<thead>
<tr>
<th>Share of YouTube Videos Worldwide that Are Flagged by Humans, by Reason for Flagging, Q4 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of total</td>
</tr>
<tr>
<td>Sexual</td>
</tr>
<tr>
<td>Spam or misleading</td>
</tr>
<tr>
<td>Hateful or abusive</td>
</tr>
<tr>
<td>Violent or repulsive</td>
</tr>
<tr>
<td>Harmful or dangerous acts</td>
</tr>
<tr>
<td>Child abuse</td>
</tr>
<tr>
<td>Promotes terrorism</td>
</tr>
</tbody>
</table>

Note: total human flagged videos=9.3 million; a single video may be flagged multiple times and for different reasons; flagging a video does not necessarily result in it being removed
Source: YouTube, “YouTube Community Guidelines Enforcement,” April 23, 2018

Over three-quarters of videos removed by YouTube in Q4 2017 were not viewed at all before being blocked, while the remainder received some number of views.

<table>
<thead>
<tr>
<th>Share of YouTube Videos Worldwide that Are Removed Before vs. After Receiving Any Views, Q4 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of total</td>
</tr>
<tr>
<td>Before</td>
</tr>
<tr>
<td>After</td>
</tr>
</tbody>
</table>

Note: total videos removed=8.3 million
Source: YouTube, “YouTube Community Guidelines Enforcement,” April 23, 2018

Although YouTube did not provide details of how this metric may have shifted over time, there are indications that the company’s machine learning processes are getting better and faster at blocking videos before they’re seen by large numbers of users. In Q1 2018, more than half of the clips taken down for depictions of violent extremism were blocked before they were seen by 10 or more viewers. By comparison, in Q1 2017 the vast majority of such videos were viewed by at least 100 people before being removed.

Industry executives are generally reacting positively to YouTube’s policing efforts, as well as the transparency the company has shown by publishing its Community Guidelines Enforcement report.
“YouTube has done a really good job of cleaning out the content that really shouldn’t be on the platform at all, or shouldn’t be carrying ads,” said Mike Henry, CEO of advertising analytics firm OpenSlate. “And I think that’s really important because the issue, as it started in February of last year, was about brand safety. While acknowledging that even their system can’t be perfect on that front, they have done all they can to clean it up.”

Others say recent steps are positive, but that YouTube needs to do more.

“It’s a good first step,” said Matthew Fanelli, senior vice president at MNI Targeted Media. “The curtain has been opened about a foot, and what people really want is to have it open 6 feet. They want to know YouTube is excluding all the explicit stuff, but they’re also saying, ‘I want a purview, and I want a look into every place that I’m going to be.’ That’s the goal of where marketers want to get to.”

Even as YouTube has grappled with brand safety challenges, the company’s business results have been healthy.

eMarketer’s March 2018 forecast of US YouTube net video ad revenues was unchanged from the previous version in September 2017. We expect the platform to reach $3.88 billion in US video advertising in 2020, up from $3.23 billion in 2018. Further, we expect video’s percentage of overall YouTube revenues to decrease slightly in that same time frame, from 73% in 2018 to 71% in 2020. This is because of the prevalence of overlays and other nonvideo formats on the service.

Although our estimates of YouTube revenues themselves haven’t changed, we made a downward revision in the company’s share of the US video ad pie. We now expect this figure to drop from 18.1% in 2018 to 16% in 2020. In our September forecast, we had projected YouTube’s market share to remain constant in the range of 20% to 22% for the entire forecast period. The reasons for the change are simple: Overall video ad spending has increased faster than we expected, and the market has grown more competitive, with more players vying for video ad dollars.
YouTube’s viewer numbers have also continued to flourish despite the company’s brand safety skirmishes. Its worldwide audience is expected to reach 1.86 billion in 2021, up from 1.58 billion in 2018. As these numbers grow by single-digit percentages in the next three years, they will also make up a slightly larger share of total video viewers, indicating that YouTube will not only grow its ranks around the world but also capture a larger market share as it goes.

### YouTube Viewers and Penetration Worldwide, 2016-2021

<table>
<thead>
<tr>
<th>Year</th>
<th>Billions</th>
<th>% Change</th>
<th>% of Digital Video Viewers</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>1.35</td>
<td>13.0%</td>
<td>66.3%</td>
</tr>
<tr>
<td>2017</td>
<td>1.47</td>
<td>9.2%</td>
<td>66.2%</td>
</tr>
<tr>
<td>2018</td>
<td>1.58</td>
<td>7.5%</td>
<td>66.4%</td>
</tr>
<tr>
<td>2019</td>
<td>1.68</td>
<td>6.6%</td>
<td>66.5%</td>
</tr>
<tr>
<td>2020</td>
<td>1.78</td>
<td>5.6%</td>
<td>66.6%</td>
</tr>
<tr>
<td>2021</td>
<td>1.86</td>
<td>4.9%</td>
<td>66.8%</td>
</tr>
</tbody>
</table>

Note: individuals of any age who watch video via YouTube app or website at least once per month
Source: eMarketer, Jan 2018

These indicators point to YouTube’s resilience in the face of the brand safety challenges it has faced since at least early 2017. Despite predictions from at least one analyst at that time that ad boycotts could cost parent company Google up to $750 million, the company has seemingly not missed a beat.

Google’s Q1 2018 earnings rose to $31.15 billion, a 25.8% increase over the Q1 2017 total of $24.75 billion. Similarly, its closing share price on May 1, 2018, was $1,060.32, which represented a 33.1% increase over the January 1, 2017, closing of $796.79. Although Google does not break out YouTube results, eMarketer estimates that YouTube’s net worldwide ad revenues totaled $7.80 billion in 2017, a 39.8% jump over the 2016 figure of $5.58 billion. These are net figures after subtracting traffic acquisition and content acquisition costs to partner sites.

### US Companies Using YouTube for Marketing Purposes, 2015-2019

<table>
<thead>
<tr>
<th>Year</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>42.7%</td>
</tr>
<tr>
<td>2016</td>
<td>45.6%</td>
</tr>
<tr>
<td>2017</td>
<td>48.2%</td>
</tr>
<tr>
<td>2018</td>
<td>50.2%</td>
</tr>
<tr>
<td>2019</td>
<td>51.9%</td>
</tr>
</tbody>
</table>

Note: companies with 100+ employees; includes use of any of the proprietary public-facing tools on YouTube for marketing purposes; includes organic and paid functions
Source: eMarketer, Nov 2017

Google’s Q1 2018 earnings rose to $31.15 billion, a 25.8% increase over the Q1 2017 total of $24.75 billion. Similarly, its closing share price on May 1, 2018, was $1,060.32, which represented a 33.1% increase over the January 1, 2017, closing of $796.79. Although Google does not break out YouTube results, eMarketer estimates that YouTube’s net worldwide ad revenues totaled $7.80 billion in 2017, a 39.8% jump over the 2016 figure of $5.58 billion. These are net figures after subtracting traffic acquisition and content acquisition costs to partner sites.


<table>
<thead>
<tr>
<th>Year</th>
<th>US</th>
<th>Non-US</th>
<th>Worldwide</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>$2.92</td>
<td>$2.66</td>
<td>$5.58</td>
</tr>
<tr>
<td>2017</td>
<td>$3.88</td>
<td>$3.93</td>
<td>$7.80</td>
</tr>
<tr>
<td>2018</td>
<td>$4.43</td>
<td>$4.70</td>
<td>$9.13</td>
</tr>
<tr>
<td>2019</td>
<td>$4.96</td>
<td>$5.54</td>
<td>$10.50</td>
</tr>
<tr>
<td>2020</td>
<td>$5.47</td>
<td>$6.29</td>
<td>$11.76</td>
</tr>
</tbody>
</table>

Note: net ad revenues after company pays traffic acquisition costs (TAC) and content acquisition costs (CAC) to partner sites; includes banners, rich media, search and video ads
Source: company reports; eMarketer, March 2018
YOUTUBE’S BRAND SAFETY CHALLENGES HAVE NOT AFFECTED BUSINESS RESULTS

GOOGLE WORLDWIDE REVENUES
Q1 2017 – Q1 2018
+25.8%

GOOGLE SHARE PRICES
Jan 1, 2017 – May 1, 2018
+33.1%

Top YouTuber PewDiePie dropped after posting anti-Semitic videos

Brands ditch YouTube after ads run next to objectionable content; analyst estimates boycott will cost Google $750 million

YouTube touts “better detection and faster removal driven by machine learning”

Some marketers suspend YouTube campaigns after ads run next to child exploitation videos; YouTube takes down hundreds of accounts and removes 150,000 videos

YouTube suspends ads on Logan Paul’s channels after vlogger posts footage of a suicide victim

Brands ditch YouTube after ads run next to objectionable content; analyst estimates boycott will cost Google $750 million

Some marketers suspend YouTube campaigns after ads run next to child exploitation videos; YouTube takes down hundreds of accounts and removes 150,000 videos

YouTube again raises video monetization thresholds in an effort to ward off “bad actors”

CNN reports that YouTube ran ads from hundreds of brands next to extremist content

YouTube raises thresholds for video monetization under YouTube Partner Program (YPP)

Credit Suisse says ad suspensions had no effect on Google’s business, and that most advertisers returned

YouTube debuts quarterly Community Guidelines Enforcement report detailing its video policing efforts

While the YouTube figures are a relatively small portion of Google’s total revenues—which include search and display advertising, and also ancillary sources such as devices—it’s notable that both YouTube and its parent company posted significant revenue increases during a year when they faced a battery of brand safety challenges.

These positive results reflect YouTube’s enduring appeal to marketers, according to Todd Krizelman, CEO and co-founder of marketing analytics firm MediaRadar.

“The moment the unsavoriness is discovered in the press, the agency that represents the brand, or the brands themselves, will immediately disavow their relationship with any broadcaster or publisher that exposed them to the liability,” Krizelman said. “What I think is interesting is how ephemeral their reaction is. If someone is doing a large campaign and it turns out someone catches unsavory video content, they will pull it, but then they’ll come back, sometimes within a week or just a few days. Very few companies say, ‘You know what? We’re really turned off by this forever and we’re just going to stop working with you.’”

To Krizelman’s point, companies that initially stepped back from YouTube but eventually returned to the platform include Procter & Gamble, Ford Motor Co., PepsiCo, Hulu and Groupon, according to an article in Business Insider on May 3, 2018.

### FACEBOOK

In May 2018, Facebook published the Community Standards Enforcement Preliminary Report, the first in what it says will be an ongoing series of reports detailing its efforts to flag and block offensive content of all types: photos, videos, images and comments. The report covered the period from October 2017 through March 2018, and future reports—which will be released periodically but not necessarily quarterly—will also allow readers to look back on earlier data for comparison.

The first study focused on six areas of violative content: graphic violence, adult nudity/sexual nudity, terrorist propaganda, hate speech, spam and fake accounts. For each of those content types, the report attempted to answer four questions:

- How prevalent are Community Standards violations on Facebook?
- How much content does the platform take action on?
- How much violating content does Facebook find before users report it?
- How quickly does the company take action on violations?

By far the most prevalent types of noncompliant content on Facebook are spam and fake accounts, which exceed other types of content by orders of magnitude. In Q1 2018, Facebook took action on 836.0 million instances of spam, up from 727.0 million in Q4 2017. Similarly, Facebook acted on 583.0 million cases of fake accounts in Q1 2018, down from 694.0 million the previous quarter.

With a time span of only two quarters, and without much context around the data, it’s difficult to put it into broader perspective and see how these types of content, and Facebook’s response, may have evolved over time. It is also unclear why cases of spam increased while fake accounts decreased in Q1 2018. But given the sheer numbers of bogus accounts, flagging and taking them down is a massive challenge for Facebook. As the study indicates, fake accounts are “often the starting point for other types of standards violations,” so keeping them in check presumably helps stem the tide of other types of problem content downstream.

<table>
<thead>
<tr>
<th>Number of Content Pieces that Violate Facebook Standards, by Violation Type, Q4 2017 &amp; Q1 2018 millions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violation Type</td>
</tr>
<tr>
<td>Fake accounts</td>
</tr>
<tr>
<td>Spam</td>
</tr>
<tr>
<td>Adult nudity and sexual activity</td>
</tr>
<tr>
<td>Graphic violence*</td>
</tr>
<tr>
<td>Hate speech**</td>
</tr>
<tr>
<td>Terrorist propaganda*</td>
</tr>
</tbody>
</table>

Note: violating content on which Facebook took action (e.g., removing a piece of content from Facebook, covering photos or videos that may be disturbing to some audiences with a warning, disabling accounts); includes posts, images, videos or comments; *increase is primarily due to improvements in detection technology; **Source: Facebook, “Community Standards Enforcement Report,” May 15, 2018

Policing Video Content on YouTube, Facebook and Twitter

©2018 EMARKETER INC. ALL RIGHTS RESERVED
The amount of content depicting adult nudity and sexual activity was constant during the study period, with 21.0 million pieces in each of the two quarters tracked. Content featuring graphic violence, terrorist propaganda and hate speech was less prevalent, with numbers in the single-digit millions, but all three of those categories saw significant increases, with graphic violence instances nearly tripling from 1.2 million in Q4 2017 to 3.4 million in Q1 2018.

Again, the data set doesn’t give enough context to draw conclusions as to why some content types fluctuate more than others. A longer look back at this data might help explain, for example, if occurrences of hate speech have increased as online discourse has grown more polarized in the US and elsewhere, but otherwise rationalizing the ups and downs in these categories is challenging at best.

What is clear from other data in the study, though, is that Facebook is making progress in some of the areas that saw the biggest increases in violative activity during the study period. In Q1 of this year, the company flagged 85.6% of graphic violence instances before they were reported by users, up from 71.6% the previous quarter. It also increased the rate at which it flagged hate speech, to 38.0% from 23.6%.

<table>
<thead>
<tr>
<th>Violation Type</th>
<th>Q4 2017</th>
<th>Q1 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spam</td>
<td>99.8%</td>
<td>99.7%</td>
</tr>
<tr>
<td>Terrorist propaganda*</td>
<td>96.9%</td>
<td>99.5%</td>
</tr>
<tr>
<td>Fake accounts</td>
<td>99.1%</td>
<td>98.5%</td>
</tr>
<tr>
<td>Adult nudity and sexual activity*</td>
<td>94.4%</td>
<td>95.8%</td>
</tr>
<tr>
<td>Graphic violence*</td>
<td>71.6%</td>
<td>85.6%</td>
</tr>
<tr>
<td>Hate speech*</td>
<td>23.6%</td>
<td>38.0%</td>
</tr>
</tbody>
</table>

Note: read as 71.6% of content that violated graphic violence standards and were acted on during Q4 2017 was flagged by Facebook before users reported it; violating content on which Facebook took action (e.g., removing a piece of content from Facebook, covering photos or videos that may be disturbing to some audiences with a warning, disabling accounts); includes posts, images, videos or comments; *increase due to improved technology and processes to find and flag content before users report it; Source: Facebook, “Community Standards Enforcement Report,” May 15, 2018.

Clearly, Facebook has more work to do to bring those figures up to the 90% range, as it has done with nudity, terrorist propaganda, spam and fake accounts. But the improvements in its early flagging of violence and hate speech indicate that its algorithms are getting better at recognizing those types of content. With hate speech in particular, even recognizing it, or allowing people to repost an instance of it to make a counterargument, can present a huge challenge.

While these figures don’t specifically relate to video, Facebook’s decision to consolidate its premium video programming under the Watch tab gives brand marketers reassurance that at least some forms of video are shielded from the types of violations that can occur more easily with user-generated content, including live streams via Facebook Live.

Facebook Watch launched in mid-2017 and is ramping up this year with entertainment shows featuring the likes of Jada Pinkett Smith, and news programming under the direction of the company’s head of news partnerships, former TV news anchor Campbell Brown.

“The more that you can curate, then the less the chances of things going wrong,” said Susan Schiekofer, chief digital investment officer at GroupM. “Facebook is leaning in a big way into Watch. They’re working with partners. They’re renewing shows. That is a more brand-safe video environment for clients.”

The challenges that Facebook has encountered since early 2017 include accusations of fake news, interference in the 2016 election, instances of violent acts streamed on Facebook Live and the Cambridge Analytica scandal. Despite these challenges, which affected average users and brand marketers alike, Facebook’s business did not suffer during this period.

Its Q1 2018 earnings totaled $11.97 billion, a 49.0% increase from the Q1 2017 figure of $8.03 billion. Its closing share price on May 1, 2018, was $185.74, up 42.5% over the $130.32 closing on January 1, 2017.

In addition to posting robust earnings and value to shareholders during the period in question, Facebook also emerged unscathed from public campaigns such as the #DeleteFacebook hashtag, which arose in the immediate aftermath of the Cambridge Analytica crisis but never gained enough momentum to make a serious dent in user numbers.

“Amidst all this scrutiny, Facebook is still doing really well,” said Serena Budhwani, account director at creative agency Sandbox. “They’ve seen ad spending go up significantly on the platform, and they’re confident enough to move forward with the release of [a] dating app. It’s not like people are leaving in droves.”

For more on Facebook’s struggles with data privacy and brand safety, see eMarketer’s June 2018 report, “Changes to Facebook Advertising After Cambridge Analytica: What’s Happening with Usage, Advertising and Data Privacy.”
FACEBOOK’S BUSINESS SOLID DESPITE FAKE NEWS, CAMBRIDGE ANALYTICA SCANDAL

FACEBOOK WORLDMIDE REVENUES
Q1 2017 – Q1 2018
+49.0%

FACEBOOK SHARE PRICES
Jan 1, 2017 – May 1, 2018
+42.5%

Facebook says it will add 3,000 workers to monitor and remove violative content

Facebook launches Watch platform for short series

The New York Times reports that Cambridge Analytica improperly used data from millions of Facebook users

Brands including Tesla and Mozilla pause their Facebook activity

Facebook says it is using artificial intelligence (AI) to flag videos that violate its policies

Facebook launches Community Standards Enforcement report

In testimony before the US Congress, Facebook CEO Mark Zuckerberg apologizes for data privacy scandal

Facebook: We can’t stop all live stream suicides

A Facebook user posts live video of a murder he committed; the company admits it needs “to do better”

Facebook announces monetization eligibility standards aimed at promoting brand safety

Twitter’s challenge with policing video is less urgent than YouTube’s or Facebook’s, because Twitter doesn’t directly monetize tweets via video ads. However, brand safety is an ongoing concern for Twitter given the plethora of trolls and bots on the platform—and the potential for users to run afoul of company guidelines for tweeting or retweeting offensive content.

In a briefing with reporters in May 2018, Twitter CEO Jack Dorsey said the company would double down on problem accounts by examining actions such as how often people tweet about accounts that don’t follow them, or whether they’ve confirmed their email address. According to coverage of the meeting in Reuters, tweets from suspect accounts will appear lower in some areas of the service, such as search results or replies, even if the tweets themselves have not been found to violate rules.

“We want to take the burden of the work off the people receiving the abuse or the harassment,” Dorsey said, acknowledging that past efforts to fight abuse “felt like whack-a-mole.”

Part of Twitter’s effort to improve brand safety involves combatting trolls “that distort and detract from the public conversation on Twitter,” according to a May 15, 2018, blog post titled “Serving Healthy Conversation.” The post indicated that fewer than 1% of Twitter accounts are reported for abuse, but that much of the reported behavior doesn’t violate company standards. Still, accounts that get flagged for violations, whether they’ve committed them or not, can have a disproportionate and negative impact on people’s experience on Twitter.

The company said it would start considering a wide variety of signals to deprioritize potentially problematic accounts before they get reported by other users.

“In our early testing in markets around the world, we’ve already seen this new approach have a positive impact, resulting in a 4% drop in abuse reports from search and 8% fewer abuse reports from conversations,” the company wrote in the blog post. “That means fewer people are seeing tweets that disrupt their experience on Twitter.”

When it comes to video content specifically, Twitter has long restricted monetization to premium content under its Amplify partner program, which it launched in 2013. The program ensures that video ads run only against vetted, professional content from partners that include top news, sports and entertainment channels, such as Bloomberg, the NBA and A&E.

The program has been a success for Twitter. Since early 2016, video has been the its largest revenue-generating ad format, and as of Q4 2017, more than half of its business came from video ads. Video was also the company’s fastest-growing ad format in Q1 2018, and it has recently announced new video content partnerships and more ways users can incorporate video into their tweets.

Despite Twitter’s success with video, some experts say the company is playing it safe by restricting monetization to professional content.

“Twitter hasn’t solved the problem. They’ve essentially just shrunk their universe to a controlled environment,” said Chris Innes, COO at ad tech firm SteelHouse. “That may work for a short period of time, but they have to build a real solution. At some point they’re going to have to open up to every other SMB [small and medium-sized business] advertiser.”

Although Amplify is theoretically open to any advertiser that abides by Twitter’s terms of service, the program caters mainly to national brands that are looking to attach themselves to premium content from publishing partners including Bloomberg Television, the Discovery Channel, the NFL and the FIFA World Cup. Advertisers that have sponsored NFL programming through Twitter Amplify include McDonald’s, Verizon and Microsoft.

Like YouTube and Facebook, Twitter has posted strong results since early 2017, indicating that its brand safety challenges and relatively conservative approach to video monetization have not hurt its bottom line or its valuation. In Q1 2018, Twitter posted earnings of $665 million, or 21.3% better than its Q1 2017 figure of $548 million. In that time frame, its share price has risen by 93.2%, closing at $34.04 on May 1, 2018, up from $17.62 on January 1, 2017.
TWITTER’S BOTTOM LINE UNHURT BY BRAND SAFETY CHALLENGES

TWITTER WORLDWIDE REVENUES
Q1 2017 – Q1 2018 +21.3%

TWITTER SHARE PRICES
Jan 1, 2017 – May 1, 2018 +93.2%

- Twitter starts testing pre-roll video ads with Dentsu, stressing its brand-safe premium video content
- Twitter suspends account of neo-Nazi site Daily Stormer after white supremacist rally in Charlottesville, Va.
- Twitter issues update of its efforts to combat misinformation in the wake of the 2016 presidential election
- Twitter launches an alpha test of a programmatic advertising and real-time bidding (RTB) platform
- Twitter says more than half its business came from video in Q4 2017
- A couple in Finland accidentally live streams their death on Twitter’s Periscope video platform
- Twitter updates its rules governing prohibited content
- President Trump retweets incendiary anti-Muslim videos from controversial group
- A federal judge rules Trump’s blocking of Twitter users who disagree with him is unconstitutional
- Twitter unveils new election security measures ahead of 2018 midterms
- In congressional testimony, Twitter says it is “determined to meet” the challenge of abuse on its platform

BRAND SAFETY: A SHARED RESPONSIBILITY

As YouTube and social media companies grapple with buttressing their platforms against abuse and other forms of problematic content, a marketer-side study illustrates the scope of the challenge from a different vantage point.

A November 2017 GumGum survey conducted by Custom indicated that two-thirds of US marketers have had their brands, or their clients’ brands, exposed to brand safety issues. For 43% of respondents, these issues have occurred a few times, while 9% have experienced them regularly and 15% once. Only 33% of participants said they’d never had these issues.

<table>
<thead>
<tr>
<th>Frequency with Which Their Brand* Has Been Exposed to Brand Safety Issues According to US Marketers, Nov 2017</th>
<th>% of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regularly</td>
<td>9%</td>
</tr>
<tr>
<td>A few times</td>
<td>43%</td>
</tr>
<tr>
<td>Once</td>
<td>15%</td>
</tr>
<tr>
<td>Never</td>
<td>33%</td>
</tr>
</tbody>
</table>

Note: includes agencies, brands and digital publishers; or their clients
Source: GumGum, “The New Brand Safety Crisis” conducted by Custom (a Digiday Media Agency), Jan 23, 2018

While it’s anyone’s guess whether the level of scrutiny that’s been placed on video and social media platforms will eventually result in more US government regulation, some industry executives anticipate a tighter regulatory framework, which could lead to more stability and trust in the platforms.

Michael Priem, CEO of Modern Impact, said: “We inevitably are going to see a bunch of new regulations, and it’s been an ongoing idea. As we’ve evolved and become more sophisticated, we need to have guidelines. I don’t think that regulations, in nature, always mean they’re going to be bad or negative.”

And even though most of the media coverage around brand safety and other issues has focused on the platforms themselves, marketers also play a part in ensuring that brands are kept safe in the digital jungle.

An April 2018 Netflix statement responding to the most recent case of YouTube inadvertently placing ads next to violative content alluded to this shared responsibility: “We employ numerous filters to avoid having our content appear on sites or videos that clearly don’t represent us or our values,” read the statement. “While that works well most of the time, there are a small number of instances where it doesn’t and we are working closely with Google to close that gap further.”

The statement refers to a joint effort by Netflix and Google and implicitly acknowledges that perfection is not a realistic goal. Rather, the companies are working to “close the gap further,” a tacit recognition that the best anyone can hope is to minimize—not eliminate—brand risk.

Andrea Ching, CMO at OpenSlate, also drove home the point about brands’ role in ensuring safety.

“It’s critically important that brands lean into understanding their own suitability guidelines and make sure they have a strategy in place that where they advertise meets their standards,” she said. “Brands and marketers have to make sure they have the types of partners that can help them deliver on those standards across all these major platforms at scale.”

KEY TAKEAWAYS FOR MARKETERS

■ The risk to brands of having their content placed next to offensive content is real and should not be minimized. Brands live and die by their reputations, so they are right to jealously guard their public image across traditional, digital and social media.

■ YouTube, Facebook and Twitter have all stepped up their policing efforts, and the communication strategies around them. These platforms will continue to hire more staff as new situations arise and, more to the point, refine their machine learning algorithms to better catch violative videos and other pieces of content before they are seen. Nevertheless, none will ever be 100% safe while they continue to host user-generated media.

■ Brand safety issues have had little or no apparent effect on top-line revenues or valuations at the three companies in question. Their business results have not seemed to suffer as a result of the many challenges they’ve faced since early 2017.
Marketers continue to see value in these platforms, which are the core of most digital video advertising strategies. For better or worse, the duopoly of Google and Facebook continues to thrive. Still, marketers need to continue to monitor usage and consumer attitudes toward these platforms, and be prepared to react quickly to any changes.

Although the onus is on media platforms to provide brand-safe environments, advertisers also have a role to play in monitoring where and how their brand messages appear.

While the business impact of brand safety issues appears to be limited so far, caution is in order for marketers. New issues have cropped up in Q2 2018 that could affect companies’ bottom lines and share prices in upcoming earnings reports, and there’s no telling what could happen if these issues accelerate or new ones arise. And even if there’s no short-term impact on business results, the cumulative effect of brand safety challenges could erode trust in platforms and drive down ad spending.

**EMARKETER INTERVIEWS**

**The Next Battle for Brand Safety Is in the Comment Section**

Susan Schiekofer
Chief Digital Investment Officer
GroupM
*Interview conducted on May 2, 2018*

**Are Brands Shifting Ad Dollars After Cambridge Analytica?**

Angela Yang
Director, Connections
T3
*Interview conducted on May 8, 2018*

Serena Budhwani
Account Director
Sandbox
*Interview conducted on May 7, 2018*

Andrea Ching
CMO
OpenSlate
*Interview conducted on May 10, 2018*

**Matthew Fanelli**
Senior Vice President
MNI Targeted Media
*Interview conducted on May 21, 2018*

**Christopher Hansen**
Chief Product Officer
IgnitionOne
*Interview conducted on May 9, 2018*

**Mike Henry**
CEO
OpenSlate
*Interview conducted on May 10, 2018*

**Chris Innes**
COO
SteelHouse
*Interview conducted on May 4, 2018*

**Todd Krizelman**
CEO
MediaRadar
*Interview conducted on May 16, 2018*

**Dino Myers-Lamptey**
UK Managing Director
MullenLowe Mediahub
*Interview conducted on May 8, 2018*

**Michael Priem**
CEO
Modern Impact
*Interview conducted on May 9, 2018*

**Tuomo Riekki**
CTO and Founder
Smartly.io
*Interview conducted on May 4, 2018*

**Victor Wong**
CEO
Thunder
*Interview conducted on May 10, 2018*
READ NEXT

Changes to Facebook Advertising After Cambridge Analytica: What's Happening with Usage, Advertising and Data Privacy
Facebook Advertising 2018: eMarketer Tackles Big Questions on Usage, Privacy and AdTargeting in Wake of Cambridge Analytica
Global Digital Video Viewers: eMarketer’s Estimates and Forecast for 2016—2021, with YouTube and Mobile Video Numbers
Q2 2018 Digital Video Trends: Monetization, Audience, Platforms and Content

SOURCES

AdExchanger
AdNews
Bloomberg
Business Insider
CNBC
CNET
CNN
Facebook
Google
GumGum
MarketWatch
New York Post
National Public Radio (NPR)
Recode
Reuters
The Daily Mail
The Hill
The New York Times
The Wall Street Journal
Twitter
US Senate
Variety
YouTube

EDITORIAL AND PRODUCTION CONTRIBUTORS

Michael Balletti  Copy Editor
Katie Hamblin  Chart Editorial Manager
Dana Hill  Director of Production
Ann Marie Kerwin  Executive Editor, Content Strategy
Stephanie Meyer  Senior Production Artist
Heather Price  Managing Editor, Content
Erica Walker  Copy Editor
Coverage of a Digital World

eMarketer data and insights address how consumers spend time and money, and what marketers are doing to reach them in today’s digital world. Get a deeper look at eMarketer coverage, including our reports, benchmarks and forecasts, and charts.

Confidence in the Numbers

Our unique approach of analyzing data from multiple research sources provides our customers with the most definitive answers available about the marketplace. Learn why.

Customer Stories

The world’s top companies across every industry look to eMarketer first for information on digital marketing, media and commerce. Read more about how our clients use eMarketer to make smarter decisions.

Your account team is here to help:

Email research_requests@emarketer.com to submit a request for research support, or contact accounts@emarketer.com or 866-345-3864 to discuss any details related to your account.